Numerical Weather Prediction

Using global increments in mesoscale analyses

Forecasting Research Technical Report No. 443

Nikos Christidis

email: nwp_publications@metoffice.com

©Crown Copyright

Forecasting Research Technical Report No. 443

Using global increments in mesoscale analyses

by

Nikos Christidis

June 7th 2004

Abstract

Better coupling between global and mesoscale analyses can improve the quality of the latter, either through additional observations, or scales that are not treated in Limited Area Models (LAMs). Global analysis increments are employed here to communicate global information into the mesoscale analyses with two different methods. One approach involves the construction of a new first guess state relative to which the analysis increments are computed. An alternative approach is to simply introduce the global increments into the analysis in the vicinity of the boundaries. Assessment of the impact of the global increments on the forecast shows that the two methods yield an opposite signal. The new first guess state leads to a detriment, whereas the boundary increments to an improvement of the forecast. The contrast is especially evident in surface temperature forecasts. This discrepancy implies that improving the analysis near the boundaries has potentially more benefit than introducing global model information in the interior domain.

Contents

1. Introduction 4

1.1 Using global increments to construct a new first guess state 5
1.2 Using global increments near the boundaries 6

2. A single observation experiment 7

2.1 Results with the new first guess state 7
2.2 Results with the boundary global increments 8

3. Numerical implementation 9

3.1 Generation of the new first guess state 9
3.2 Limitations of the new first guess approach 11
3.3 Boundary global increments in LAM analyses 13

4. Results 14

4.1 Case Studies 14
4.2 Real time trial 15

5. Conclusions 17

Appendix: Including the global increments in 3DVAR runs 19

Figures 20

References 32

1. Introduction

Discrepancies between global and mesoscale analyses valid at the same time may arise due to a) assimilation of different observations over the common area, b) information from observations outside the LAM domain over the common area which is absent in the mesoscale analyses, c) difference in the number of modes used in horizontal transforms to represent the analysis increments and d) lack of information from unresolved scales in mesoscale analyses (i.e. the problem of representing the harmonic wind discussed later). The motivation for the development of the two schemes discussed in this report is to improve the coupling of global and mesoscale analyses by introducing information from the global model into the mesoscale 3DVAR system (*Lorenc et al.* 2000). Global increments are used to either construct a model state to which the mesoscale analysis increments will be added, or only to correct the mesoscale analysis increments near the model boundaries.

Schemes that make use of global increments have the following advantages:

- They ensure a smoother transition between global and mesoscale analyses, as the analysis increments at the boundaries reduce to zero and the analysis there will be determined by the global increments.

- The global increments communicate information from observations outside the mesoscale domain, but also from those inside and near the boundaries. Much of this information from observations in the interior is absent from the mesoscale analysis increments, as they reduce to zero at the boundaries.

- The construction of a new first guess also communicates information inside the LAM domain through modes that are not resolved in the transforms that produce the analysis increments. This information may provide a way of representing the harmonic wind component (*Lynch*, 1988), which corresponds to large scale flow and cannot be identified as being either divergent or rotational. An additional term to the rotational and divergent contribution in the kinetic energy is required to adequately describe the flow in LAMs. This term is referred to as harmonic wind and its contribution has so far been omitted in mesoscale analyses.

- The use of global increments only near the boundaries is a simple approach that does not interfere with the minimisation of the cost function and can be easily incorporated into future developments (4DVAR). On the other hand, the new first guess approach enables the use of a new guess state, different from the background, a feature that has been present but not yet functional in the context of mesoscale 3DVAR.

The two approaches of introducing global information into the mesoscale analysis are described in sections 1.1 and 1.2. A simple example from a single observation experiment is presented in section 2. The implementation of the methods in the 3DVAR system and some issues that need to be considered when the global

increments are employed are discussed in section 3. Results from a set of case studies and from a real time trial are given in section 4 and the main conclusions from the project are summarised in section 5.

1.1 Using global increments to construct a new first guess state

The mesoscale 3DVAR system computes analysis increments of a model perturbation, $\delta\mathbf{x}$, relative to a first guess state, $\mathbf{x}_g$, that in most applications is identical to a prior (background) estimate, $\mathbf{x}_b$, from a T+3 forecast. It is called "first" guess, as the guess solution is updated in the iterative process of computing the analysis increments, which are still relative to the first guess. 3DVAR may run at a lower resolution than the forecast model in order to reduce the cost, in which case the state vector at low resolution, $\mathbf{w}$, is transformed to model resolution with the simplification operator, S, and it is therefore $\mathbf{w} = S(\mathbf{x})$.

The analysis should in principle be independent of the first guess state, which, however, is not the case for two reasons. First, the analysis increments are constructed as the composite of a finite number of horizontal modes and therefore components that are omitted in the horizontal transforms enter the analysis from the state onto which the increments are added (background, or here, the new first guess state). It is these components that may convey useful information about the harmonic wind. Second, the analysis increments near the boundaries gradually reduce to zero through a double sine function and the analysis there comes from the guess state. The common choice of the background state does not communicate information from the global analysis, which could prove very useful, as the presence of observations either side of the boundaries is taken into account and inconsistencies between the global and the mesoscale analysis reduce.

The new first guess state is produced by adding global increments to the background state

$$\mathbf{x}_g = \mathbf{x}_b + \delta\mathbf{x}_{GL} \,, \qquad (1)$$

where the global increments, $\delta\mathbf{x}_{GL,}$ are defined as the difference between a prior (background) estimate from a T+6 run of the global model, $\mathbf{x}_{b,GL}$, and the global analysis, $\mathbf{x}_{an,GL}$, both reconfigured to the mesoscale domain

$$\delta\mathbf{x}_{GL} = \mathbf{x}_{an,GL} - \mathbf{x}_{b,GL} \,. \quad (2)$$

The new first guess in Eq. (1) includes both the effect of small scale features (term $\mathbf{x}_b$) and also information from the global analysis (term $\delta\mathbf{x}_{GL}$) and hence can provide a smooth transition between the mesoscale and global analyses near the boundaries. The presence of the term $\delta\mathbf{x}_{GL}$ may also provide a representation of the harmonic wind in the LAM domain. The analysis increments minimise the penalty function:

$$\mathrm{J}(\delta\mathbf{w}) = \frac{1}{2}(\delta\mathbf{w}_b - \delta\mathbf{w})^T \mathbf{B}^{-1}(\delta\mathbf{w}_b - \delta\mathbf{w}) + \frac{1}{2}(\mathbf{y}_o - \mathbf{y})^T \mathbf{R}^{-1}(\mathbf{y}_o - \mathbf{y}) \,, \qquad (3)$$

where $\mathbf{B}$ and $\mathbf{R}$ are the background and the observational error covariance matrices respectively, $\mathbf{y}_o$ the observations and $\mathbf{y}$ the model predicted observations. Following the conventional notation (*Ide et al.*, 1997), the delta (δ) quantities denote perturbations with regard to the FG state.

For the current choice of the background perturbation, $\delta\mathbf{w}_b$ is zero (identical background and first guess), but for the new choice it will be:

$$\delta\mathbf{w}_b = S(\mathbf{x}_b) - S(\mathbf{x}_g) = \mathbf{w}_b - \mathbf{w}_g \ , \qquad (4)$$

which is equal to minus the global increments on the 3DVAR resolution (Eq. (1) and (4)). The analysis is given by

$$\mathbf{x}_a = \mathbf{x}_g + S^{-1}(\delta\mathbf{w}) . \qquad (5)$$

Since analysis increments, $\delta\mathbf{w}$, will reduce to zero at the boundaries relative to the first guess, the mesoscale analysis at the boundaries will match the global analysis.

Dealing with inconsistencies between the global and mesoscale analyses, rises as an important issue in the effort to optimise the initial conditions for limited area models. In a different approach, *Brozkova et al.* (2001) employed the digital blending method to combine a large scale analysis with a mesoscale forecast. They kept the shortwave modes of model fields from the mesoscale forecast and the longwave ones from the global analysis and determined a transition zone of blending global and mesoscale modes, by taking into account differences in resolution. A digital filter blending method was employed. This blending of long and short wavelengths from a global and a mesoscale model is another way of utilising all the available information in the construction of initial conditions, designed within the context of spectral models. However, the shortwave component was derived from a mesoscale forecast rather than an analysis, hence all the information from observations comes from the longwave component. A more appropriate blending of global and mesoscale analyses with the same technique is introduced in *Siroka et al.* (2001). Unlike the blending method, which is performed externally, in the approach followed here, the global increments are incorporated into the 3DVAR system.

1.2 Using global increments near the boundaries

An alternative approach to introducing a new first guess state is to limit the use of the global increments only in the vicinity of the boundaries. This method concentrates on improving the analysis near the boundaries by making better use of the observations in that area, without allowing any information from the global model in the interior of the LAM domain. The better use of observations around the boundaries is expected to have a positive impact on the quality of the mesoscale analysis, whereas the benefit from global information in the inner model domain is more speculative and will be assessed with the new first guess state approach.

An important advantage of the boundary global increments scheme is its simplicity. It does not interfere with the minimisation of the cost function, but allows 3DVAR to run as usual with the background as the first guess. At the end of the minimization an increment blending takes place, whereby the global increments are combined with the analysis increments with a scaling factor that reduces from 1 close to the boundaries (only global increments) to 0 further away into the LAM domain (only analysis increments). Because the scheme does not interfere with 3DVAR

computations, it also manages to overcome a number of problems that arise with the previous method, as will be discussed in the section 3.3.

The global increments have so far been defined as the difference between the global analysis and the global background, i.e. as the global model analysis increments. However, analysis increments from the global and the mesoscale model were found to be dissimilar (Fig. 1), both in terms of their features and of their magnitude. Therefore, defining the global increments as the difference between the global analysis (reconfigured to the mesoscale domain) and the mesoscale background might be a better choice. The disadvantage of this new definition is that the two models show differences in orography, as they run on different resolution and so subtracting the two fields will reflect some of the orographic discrepancy, especially at lower levels. The original definition may not be quantitatively accurate, but it is still indicative of the sign and the relative impact of observations around the boundaries and may thus be an adequate correction to the mesoscale analysis. Both definitions will be tested for the case of the boundary global increments, but only the original definition will be considered in the new first guess method.

2. A single observation experiment

This section shows results from a simple, idealised, single-observation experiment. A unit size U-wind anomaly at model level 1 is specified as a single observation and is placed at either side of the western boundary of the LAM domain. The purpose of the experiment is to illustrate the extra information that the global increments communicate into the mesoscale analysis.

2.1 Results with the new first guess state

Analysis increments from a single observation experiment with the new first guess state are shown in Fig. 2. When the observation is located outside the model domain (left plot) and there are no other observations to assimilate, the analysis increments relative to the background should be zero, or, in other words, 3DVAR is in this case trying to reproduce the background. Combining Eq. (1) and (5), the analysis can be written as:

$$\mathbf{x}_a = \mathbf{x}_b + \delta\mathbf{x}_{GL} + S^{-1}(\delta\mathbf{w}) . \qquad (6)$$

Nevertheless, the analysis increments relative to the first guess, $S^{-1}(\delta\mathbf{w})$, are not exactly equal to $-\delta\mathbf{x}_{GL}$, as they are constructed with only a limited number of modes in the horizontal transforms. Modes present in the global increments, which are not treated by the horizontal transforms will also be present in the analysis increments relative to the background ($\delta\mathbf{x}_{GL} + S^{-1}(\delta\mathbf{w})$). This explains the non zero analysis increments in Fig. 2, which comprise the information from the global model that goes into the mesoscale analysis. The higher values of the increments near the western boundary are a result of the double sine function, which reduces term $S^{-1}(\delta\mathbf{w})$ and thus allows more of the global increments to enter the analysis. The maximum value of the global increments over the LAM domain is 0.195 (see Fig. 3)

and the final analysis increments with the new first guess communicate approximately up to half of this value.

When the observation is located within the model domain (right plot) the analysis increments relative to the background are of course non zero and should represent the distribution of the impact of the observation in the domain. Horizontal modes not treated by the transforms are in this case more evident away from the observation, where the impact of the observation is small. Near the boundaries the global increments enter again the analysis through the double sine function and thus preserve some information from the observation.

2.2 Results with the boundary global increments

The same experiment was carried out with the boundary global increments method. In this case 3DVAR calculates the analysis increments relative to the background and subsequently blends the global increments with the analysis increments in a rim of 9 grid points in the boundary area (see section 3.3).

Fig. 3 depicts the analysis increments from the global and the mesoscale models with the observation outside the model domain. The mesoscale analysis contains information from the observation in the vicinity of the boundaries and shows the greatest impact at the nearest (for this case western) boundary. It should be noted that the maximum increment value in the mesoscale analysis is close to the one from the global analysis in the same domain, which means that the maximum impact of the observation is communicated into the mesoscale analysis. The analysis increments in the inner domain are zero, although the global model has non zero increments. However, in real cases with various observations, the analysis will be dominated by the observations in the inner domain, whereas in the region of the boundaries, the scheme will introduce a zone where the analysis will be determined mainly by the global model.

Fig. 4 shows the analysis increments from the global and the mesoscale model with the observation inside the model domain. Mesoscale model results are shown with and without the global increments. The analysis without the global increments reduces to zero at the boundaries and includes areas of positive (though small) increments north and south of the observation. Introducing the global increments into the analysis, has the effect of increasing the impact of the observation at the boundaries, which makes up to some extent for the effect of the double sine function and also confines the secondary positive increments to smaller areas. Again, the assimilation of more observations inside the domain will significantly change the analysis, but the increments at the boundaries will be determined mainly by the global model corrections. One point to stress is that the use of the global increments does not produce exactly the global analysis in the vicinity of the boundaries, as the global and mesoscale analysis increments are relative to different backgrounds. However, the global increments communicate the sign of the impact, as well as the relative impact of observations near the boundaries assimilated by the global model into the mesoscale analysis and act as a correction near the boundaries. How much of this correction will actually be introduced into the

forecast depends on the extent of the boundary conditions (discussion about rim width in section 3.2).

3. Numerical implementation

3.1 Generation of the new first guess state

The numerical implementation of the scheme must be designed to be computationally inexpensive, as well as efficient in terms of coding. The global increments appear in Eq. (3) at the beginning of the minimisation in terms $\delta\mathbf{w}_b = -\delta\mathbf{w}_{GL}$ and $\mathbf{y}$. Introducing term $\delta\mathbf{w}_b$ is straightforward, as 3DVAR already provides the option of using a first guess state different from the background. The main difficulty is creating the C_x columns (guess state values at the location of the observations), an operation performed by the OPS system, which requires not only the main prognostic variables, but also derived ones. Several different methods are discussed here, based on ideas by A. Lorenc (pers. comm.), originally proposed for the implementation of an outer loop in 4DVAR.

Method 1

This method utilises the Incremental Analysis Update (IAU) code to add the global increments to the model state and derive the first guess from model dumps, which are used as input to the OPS. This method requires two model runs: one to output the background state, $\mathbf{x}_b$, and one which adds the global increments and outputs the new first guess, $\mathbf{x}_g$. The second run, illustrated in Fig. 5, adds the increments before the observation window, as the new first guess is required at the observation times. The two main disadvantages of this method is a) the need of a second model run which is computationally expensive and b) the fact that the increments propagate information from the global model incorrectly, as they are added before the actual analysis time.

Method 2

This is a variant of method 1 that computes the guess state at the appropriate time, without the disadvantage of developing information from the global model in the course of the forecast. The approach would be to add the increments using the IAU code at three successive model dump times, before, at and after the analysis time and perform single timesteps to calculate the derived fields. Interpolation between the fields at the three different times would provide the first guess at the observation times. However, the method is still expensive and also the consecutive addition of the increments would give undesired shocks to the model.

Method 3

This method adds the global increments to all fields within 3DVAR when the $C\mathrm{x}$ columns are calculated. Although it avoids the need for an extra model run, it requires some software to perform the addition of the global increments at the

observation locations to the $C\mathrm{x}$ columns for the background state. The output will be the new $C\mathrm{x}$ columns for the new first guess. There is already an algorithm in 3DVAR for the updating of the $C\mathrm{x}$ columns at the end of each iteration, which can be used in the implementation of the method.

Method 4

While the previous method constructs the $C\mathrm{x}$ columns for the new first guess state before the minimisation of the cost function, an alternative way would be to add the global increments to the computed analysis increments, $\delta\mathbf{w}$, at the end of each iteration. The $C\mathrm{x}$ columns are updated on each loop by adding the first guess state at the observation locations to the new guess increments at the observation locations:

$$C\mathbf{x}^{+} = C\mathbf{x} + CS^{-1}\delta\mathbf{w}\,. \qquad (7)$$

After adding the global increments to the analysis increments, Eq. (7) becomes:

$$C\mathbf{x}^{+} = C\mathbf{x} + CS^{-1}(\delta\mathbf{w} + \delta\mathbf{w}_{GL}) = (C\mathbf{x} + CS^{-1}\delta\mathbf{w}_{GL}) + CS^{-1}\delta\mathbf{w} \qquad (8)$$

where the term $C\mathbf{x} + CS^{-1}\delta\mathbf{w}_{GL}$ in Eq. (8) is the $C\mathrm{x}$ columns for the new first guess. Although this approach is more expensive than the straightforward method 3, it is more efficient in terms of coding and easier to maintain. Cloud fractions, or other parameters that may be assimilated in the future, can also be treated without the need to generate new code for the construction of the new first guess.

From the different approaches to the generation of the new first guess, the last method seems to make the best candidate and was the one that was employed in the project. The initial implementation of the scheme in the 3DVAR system can be summarised in the following 4 stages:

1. Reading in the global increments and storing them in a new model variable. The increments are introduced via the non zero *PFBACK* option in *VarProg_AnalysePF* and are assigned to a new variable *PF_gl*. Two new subroutines (*Var_ModelAllocateCopy* and *Var_ModelCopy*) were written to copy variable *PF* (initially read through *PFBACK* in 3DVAR) into *PF_gl*.

2. Transforming the background perturbation, $\delta\mathbf{w}_{\mathrm{b}}$, into control variables, $\mathbf{v}_{\mathrm{b}}$. This involves specifying *PF_gl* when subroutine *Var_Ttransform* is called.

3. Adding the global increments to the analysis increments after each iteration. The already existing subroutine *Var_AddFracModel* is used to add *PF_gl* to the guess solution that comes out of the U transform in *Var_TotalPenAndGrad*.

4. Adding the global increments to the analysis increments at the end of the iteration process in *VarProg_AnalysePF*, to make the analysis increments relative to the background. Again, subroutine *Var_AddFracModel* is employed for this operation.

The above changes were introduced in Var stable build 19.1. Further code changes were made in the course of the project, in order to remove high frequency terms in

the horizontal transforms (see section 3.2) and to allow the aerosol field to be set up in *PF_gl*, which is required for visibility assimilation.

3.2 Limitations of the new first guess state approach

Generating the global increments

In order to construct the global increments, global model dumps at T+0 and T+6 are required. Reconfiguring the models dumps to the LAM domain and taking their difference would produce the global increments for 3DVAR. However, in an operational context there are no global analysis files (T+0) available at 03, 09, 15 and 21Z. For these runs the global increments need to be defined as the difference between T+3 and T+9 forecasts, as the nearest approximation to the actual analysis and background. The T+9 files are currently not available from the operational suite.

Filtering

When the new first guess state is used, the initial value of the cost function increases by several orders of magnitude, which is mainly the contribution of the background term (Fig. 6). Non zero increments near the boundaries are expected to increase the cost function as the background penalty will start from a non zero value. The magnitude of the increase is determined by high frequency terms generated by the horizontal transforms, associated with large coefficients. When filtering of the high frequency terms is applied to the global increments, the initial cost function decreases with respect to the number of modes retained. In many cases, like the one in Fig.6, when no filtering is applied the minimisation does not converge and eventually fails. Little filtering is also not efficient, as the number of iterations remains high and may increase even beyond the commonly adopted limit of 200. On the other hand, too much filtering may mask useful information. The *quarter* resolution (36 modes in the E-W direction and 45 modes in the N-S) is chosen as the best compromise and is used in the work presented in this report.

The maximum number of modes kept in the decomposition of the global increments is specified with two new variables *WaveMaxLAM1* and *WaveMaxLAM2* in the transform namelist. This information is then passed during the horizontal transforms to subroutine *Var_ThLAMFFT*.

Convergence

The number of iterations in the minimisation of the cost function increases when the new first guess in employed and also when visibility assimilation is included. When both options are switched on, they seem to act synergistically to increase the number of iterations in a non linear way, as illustrated in Fig. 7.

The global modes that are introduced into the analysis with the new first guess state can only be identified when the cost function has converged well to its minimum value (within a percent). The common convergence criterion in 3DVAR (*J17abs*=75) terminates the minimisation process while the cost function is still several percent

higher than its minimum, whereas the use of global increments increases the number of iterations and therefore results in better convergence. This can be misleading when comparing results from a control run without global increments and a test run with the new first guess state. The analysis from the test run gives a better fit to observation (Fig. 8a), but this is due to the better convergence rather than any benefit from the global increments. To ensure good convergence and also to allow a meaningful comparison between control and test runs, all experiments for the assessment of the new first guess method that will be presented later were run with 200 iterations. When both control and test are run with 200 iterations, the benefit from the new first guess is no longer evident in the analysis fit to observations (Fig. 8b). One question that arises is whether any potential benefit from the method will be more important than the improvement due to better convergence with a high number of iterations and, therefore, whether it can justify the extra computational cost.

One way of speeding up convergence in cases with visibility assimilation, might be to use "dry" global increments. The rationale is that as the global model does not assimilate surface humidity, the new scheme may try to fit global information that is not useful and may unnecessarily slow down the minimisation. A test experiment with dry increments showed only a small decrease in the number of iterations (120 instead of 125) and no evident improvement in the fit to humidity observations and so the original definition of the global increments was kept. Even if the dry increments managed to accelerate convergence significantly, that would only have an effect when the common convergence criteria are employed. However, such criteria are inadequate, if any impact of the new approach is to be manifest, as the final cost function must be a good representation of its minimum. This stresses the need for a high number of iterations (i.e. 200, as employed here), not only in cases with visibility assimilation, but whenever the new first guess state is used.

Visibility and PE configuration

Visibility assimilation is associated with a dependency on the configuration of the Processing Elements (PEs), which generates discrepancies when the number of PEs changes. Sensitivity studies with Var Stable Build 19.1 show that the discrepancies remain small (of the order of 10^{-3}%) in runs without the new first guess state, but become large (several tens of percent) when the global increments are introduced (Fig. 9). The problem arises only when a non-linear visibility operator is used (either fully non-linear, or with the double-loop option). Some improvement is noticeable in runs with Var Stable Build 19.4, which might be related to the correction of LAM boundary haloes. In this case the PE dependency appears only with the fully non-linear visibility operator and the discrepancy is again of the same order. However, the more recent 20.1 version, running on the SX6 platform, shows a similar dependency as Stable Build 19.1, with discrepancies arising when the assimilation is performed with 4 and 2 PEs and both non-linear versions of the visibility operator. Fig. 9 illustrates the problem from a case study with version 20.1. The configuration dependency is also likely to affect ozone assimilation, as the ozone code was based on the visibility code. Some ozone assimilation tests (D. Jackson, pers. comm.) indeed displayed discrepancies with different number of PEs.

One reason that may result in such PE dependencies is pointed out in VTDP20 and is related to the descent algorithm, which runs unparallelised. The descent algorithm was modified to allow for parallel data (which mainly involves computations of inner products) and that can introduce some dependency on configuration. It may as well be that the use of a different first guess enhances the differences in inner product computations and makes them appear earlier in the descent, though this doesn't explain why the linear version of the visibility operator does not produce a similar behaviour.

The extended width rim

Noise in the pmsl field with New Dynamics (ND) which appears as small scale waves propagating in from the boundaries and also results in a 10-20% detriment in objective verification, led to the use of an extended width rim from 4 to 8 grid points (*Wilson et al.*, 2004) as a temporary solution. The extended rim hinders much of the potential benefit from the new first guess. Until the 4 point rim is re-introduced, neither better consistency between global and mesoscale analyses, nor better use of observations around the boundaries can be demonstrated. Hence, the assessment of the scheme with model runs which use the extended width rim will only test the impact of extra modes within the LAM domain on the forecast.

3.3 Boundary global increments in LAM analyses

Fig. 10 shows a schematic illustration of the alternative method which introduces the global increments only in the vicinity of the boundaries. The global increments are combined with the final analysis increments with a scaling factor of 1, up to five grid points away from the boundaries, to 0, nine grid points away. This means that in the first 5 grid points the analysis is determined solely by the global increments, from 9 points and beyond it is determined solely from the 3DVAR increments and in the intermediate points a blending of the two is performed.

In terms of coding, the global increments are read in 3DVAR the same way as in the new first guess method and are again stored in variable *PF_gl*. New software is required to apply scaling to the fields of a PF structure. Subroutine *Var_ModelBoundaries*, was written for this purpose. As the scaling is a locally dependent operation, the development of the new subroutine had to comply with the requirements of parallel processing. The blending of the global and the analysis increments is performed in *VarProg_AnalysePF* at the end of the minimisation process, using subroutines *Var_ModelBoundaries* and *Var_AddFracModel*.

The use of boundary increments is not constrained by some of the limitations of the new first guess method. Firstly, as 3DVAR continues to use the background state as its first guess, the initial background penalty remains zero and there is no increase in the number of iterations required for convergence. The global increments are only introduced into the final solution and so there is no need for their decomposition and filtering in this case. Moreover, as the potential benefit of the method is unrelated to global modes inside the LAM domain, there is no need to run the

minimisation with a high number of iterations, in order to enhance the impact of such modes on the analysis. Of course the absence of these global modes could be a disadvantage of this method, provided that they indeed communicate useful information to the analysis. Finally, the problem with the PE dependency has a lesser effect (same as the control for the example in Fig. 9) and even though it may still be present, it generates hardly any discrepancy in the results from runs with different configurations.

The use of the extended width rim is undesirable for this method, as it masks almost all the improvement that the global increments are trying to achieve near the boundaries. For this reason, all the tests in this work employ the 4 point rim when the boundary increments are used. The testing of the new first guess method, however, was carried out with the extended rim. This implies that the new first guess method investigates the impact of global information inside the domain, whereas the boundary increments method investigates the impact of global information near the boundaries.

4. Results

A twofold approach was adopted to test the impact of the global increments on the forecast. A first assessment used a set of 6 case studies to run a control experiment without the global increments and a test experiments with the two methods of introducing them into the analysis. Case studies are useful to point out problems with the implementation of a new scheme and to assess the magnitude of the impact. However, a limited number of cases can only give a first indication of the sign of the impact, since the actual signal is small and would require a large number of cases to be clearly detected. For this reason, the assessment was also complemented with a real time trial that ran for over a month and so it includes a sufficient number of forecasts to better indicate the sign of the impact from the global increments.

4.1 Case studies

The six case studies used in the assessment cover the period April-October 2002. A suite is set up for each case with a first run at 00Z, followed by four assimilation cycles at 03, 06, 09 and 12Z. A 36hr forecast starting at 12Z provides the statistics for the objective verification. The experiments are carried out without visibility assimilation, in order to demonstrate any benefit from the global increments, which might otherwise be obscured because of the PE configuration dependency, especially so for the case of the new first guess state.

Figures 11 and 12 plot differences in the analysis with and without the global increments and show how the global increments change the analysis. Fig.11a is essentially an illustration of global modes that are not treated in 3DVAR, either due to filtering, or due to scales not represented in the mesoscale model, that enter the analysis through the new first guess. These features, initially small, propagate to the next analysis, when a new set of filtered modes is added to them and the same

procedure is repeated until the last cycle, depicted in Fig. 11b. It is interesting to note that the magnitude of the differences in the last cycle has grown to become the same as the magnitude of the analysis increments (see for example Fig. 1a), which means that the use of the new first guess state changes the final analysis increments significantly. A similar illustration for the boundary increments method is given in Fig. 12. At the first cycle (Fig. 12a) the analysis difference shows only the global increments that are introduced in the vicinity of the boundaries. This extra information propagates to the next cycle when new increments are introduced near the boundaries and the same is repeated until the final cycle, shown in Fig. 12b. Again the differences are of the same order as the analysis increments.

Probably the most useful test on the impact of the global increments comes from the objective verification of the forecast against observations. Figures 13, 14 and 15 show the change in the surface temperature and surface wind RMS error for a test experiment with the global increments relative to a control without. The results are plotted for experiments with the new first guess (Fig. 13) and the boundary increments, either as defined originally (Fig. 14), or as the difference between the global analysis and the mesoscale background (Fig. 15). On the basis of the limited number of cases, it appears that there is only a small impact of uncertain sign on temperature when the new first guess is employed, whereas the sign of the impact appears more clearly to be negative in the wind forecast. The boundary increments, as defined originally, lead to a more pronounced impact, which shows as an improvement in both temperature and wind. Figures 13 and 14 show that although the boundary increments give overall better results than the new first guess state, case study 29-Jul-02 favours the use of the new first guess. The most interesting of the six cases seems to be the 01-Oct-02, where all methods have a significant impact on the forecast. The new first guess shows a mixed effect, whereas the boundary increments lead to a significant improvement relative to the control. This case suggests that the modes of the global increments introduced inside the LAM domain bring considerably less benefit to the forecast than the information introduced by the same increments near the boundaries. The definition of the increments relative to the mesoscale background gives like the original definition a more pronounced impact than the new first guess, but the signal in this case is mixed and this suggests that the original definition is preferable. For this reason, only global increments constructed relative to the global background are employed in the real time trial investigations.

4.2 Real time trial

Results from the case studies point out that the boundary increments are more likely to improve the forecast than the new first guess. An extended trial is employed to confirm this indication and to provide a more rigorous assessment of the impact of the global increments, especially with respect to its sign. Extended trials also provide a sufficient number of forecasts to compute categorical statistics for visibility, precipitation and cloud amount. Such computations would be unreliable with only the 6 forecasts from the case studies. Both a control and a test suite are set up, to ensure that the only difference comes from the global increments. The trial periods are:

05 Dec 2003 to 14 Jan 2004, for the new first guess method and
29 Oct 2003 to 30 Nov 2003, for the boundary increments as originally defined.

Although the new first guess trial period is longer, it contains two short void periods and so the two trials actually produce the same number of forecasts. All trials include visibility assimilation, as they are designed to be a good representation of the operational suite, in order to test whether the global increments would be suitable for operational implementation. It is reminded that the assimilation with the new first guess is performed with 200 iterations, to enhance the impact of the method, and that in the boundary increments trial a 4 point rim is employed.

Figures 16 and 17 show the day to day change in the cumulative skill score (test vs. control) for visibility, precipitation accumulation, cloud amount, temperature, wind and the total weighted score for observations over the whole LAM area (block 00) and for selected stations in the LAM domain (block 03). The use of the new first guess leads to a decrease in skill for all variables, which is worse for temperature and visibility. After the early part of the trial, the skill generally levels off without any significant trend, with the exception of temperature, where there is a stepwise further decrease in the later part of the trial. In contrast, the boundary increments give a positive, though small and close to neutral signal for all parameters, which becomes significant for temperature in the later part of the trial. The increase in the skill of the temperature forecast is clearly reflected in the total weighted score.

Fig. 18 gives an illustration of the statistics for the composite of all the forecasts in the trial runs. The results show that the two methods have an opposite effect on all the parameters, with the exception of visibility in block 03, where there is a detriment in both cases. Both methods have a significant impact on temperature, while the new first guess is also associated with a large detriment in visibility. The real time trial results are consistent with those from the less stringent testing with the case studies that indicate a positive impact on temperature and wind forecasts with the boundary increments and a negative impact with the new first guess. However, the variability in the cumulative skill in the early part of the trial (Fig. 16 and 17) makes clear that conclusions from a small number of case studies may be misleading, whereas increasing the number of cases would gradually increase confidence in the results. A measure of the impact of the global increments on a larger scale is given by the mean sea level pressure rms error. Figures 19 and 20 illustrate the mean and rms error for the new first guess and the boundary global increments for the composite of the trial forecasts. The impact is small early in the forecast, but increases with forecast range and has an opposite sign for the two methods. As with the other fields, the positive impact is associated with the boundary global increments.

5. Conclusions

Introducing information from the global model into mesoscale analyses has been investigated with two different methods. The first uses a new first guess state to assess the impact of modes in the interior domain that cannot be treated by the LAM. The second uses global increments only near the boundaries to make the most of information from observations around the boundaries that is missing from the mesoscale analysis. The results presented here indicate that the latter has more potential to improve the forecast, and in particular the surface temperature.

The use of the new first guess state appears to lead to a forecast detriment. As the tests for this method are run with the extended width rim, it is the impact of global modes in the interior that is investigated. The argument that these modes carry useful information about the harmonic wind becomes difficult to justify, when considering the detriment in the wind forecast. A more conclusive investigation would also examine higher levels in the atmosphere (i.e. the jet level), where the orography has less influence. However, as the information that comes into the analysis is associated with high frequency terms rather than the large scale flow, it is doubtful whether much benefit can be achieved. The global information carried into the analysis appears in the final analysis increments as periodic wavelike features, which demand a large number of iterations to become evident. An apparent improvement in the analysis fit to observations is attributed to the better convergence rather than the global information itself and the increase in computational cost is not justified by an accompanying improvement in the forecast quality.

The idea behind the new first guess state method was that modes capable of being represented by the mesoscale control variables would be re-analysed, while modes which were not represented would be retained from the first guess. It was mistakenly thought that the represented modes would be small scale (up to the longest wavelength in the spectral control variable), and the unrepresented modes larger scale. The results in this study show this is not the case. The mistake comes from thinking spectrally when the control variable modes are multiples of half-waves, rather than waves. The modes which are not full waves are not orthogonal over the model area to larger scale waves, or indeed to a wavenumber zero mean field. This means that the mesoscale control variables are capable of representing most of the guess field – even if in reality the guess is made up of large scale patterns. Only aspects near the boundary cannot be represented (Fig. 3).

Modifying the method to allow large scale flow information to come from the global increments, could be of great benefit in high resolution modelling, where a smaller model area would not manage to represent adequately the larger scales. An alternative suggestion is being followed up by researchers in JCMM, which attempts a more explicit scale separation. The new approach considers a cut-off wavelength, below which the mesoscale analysis is to add detail, but above which the large scale guess is retained. This is achieved by filtering the global increments on their larger domain, either spectrally, or using a grid point digital filter (as tried by Meteo-

France). The predetermined increments, which can be non zero at the boundary, are added to the mesoscale analysis increments coming from spectral control variables, which are filtered in a complementary way.

Introducing global increments near the boundaries is found to improve the forecast, when the increments are defined relative to the global background. It may be expected that global increments relative to the mesoscale background would be a better definition, but the present results do not support this assumption. This indicates that differences in orography between the two models are important and that the global increments ought to be independent of the orography. The benefit of this method comes from observations near the boundaries that the mesoscale analysis does not account for. It is interesting to note that most of the improvement from the global increments appears in surface temperature, even though the global model does not assimilate surface temperature observations. If that was the case, a further forecast improvement would be expected. Finally, the positive impact on the early precipitation forecast, even though small, may still contribute to the ongoing effort of improving rainfall forecasts (*Christidis and Macpherson*, 2004). So far the method has been tested with a trial during autumn. Complementary trials in other seasons with different weather conditions would increase confidence in the benefit from the boundary increments. Operational implementation of the method would require the restoration of a 4 point rim.

Acknowledgements: The new first guess state method was proposed by A. Lorenc. I want to thank Andrew for useful discussions during this project that helped understand how to interpret the results and how to develop the original idea further. I also want to thank R. Renshaw and A. Clayton for helping with technical issues and B. Macpherson for supervising the work.

Appendix: Including the global increments in 3DVAR runs

1) Standalone runs with the varui

• Additional line in the Job Environment / Extra commands before:
`export VAR_GLB_INCR=true` for the new first guess,
or
`export VAR_GLB_INCR_BND=true` for the boundary increments.

• Global increments input file, to be specified in:
Analyse PF Program / Perturbation forecast (VAR_PFBACK)

• Additional lines in Transform namelist part 1 / Extra lines for namelist
to specify the number of horizontal modes in the x and y direction:
`WaveMaxLAM1=36`
`WaveMaxLAM2=45`

2) Suite jobs

• Additional lines in the analysis task (Extra lines before task)
`export VAR_GLB_INCR=true` for the new first guess,
or
`export VAR_GLB_INCR_BND=true` for the boundary increments.
`export VAR_PFBACK=global_inc_file`

• Var components, additional lines in file VarList_AnalysePF_ms in &Var_TransformNL:
`WaveMaxLAM1=36`
`WaveMaxLAM2=45`

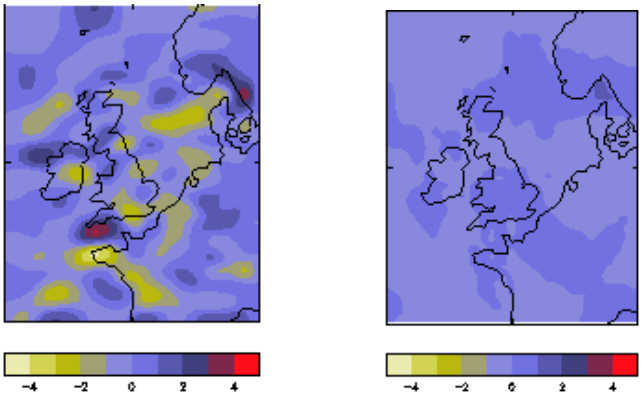

Fig. 1 The U-wind analysis increments at model level 1 from the mesoscale model (left) and the global model (right). Case: 01 Oct 2002 / 12Z

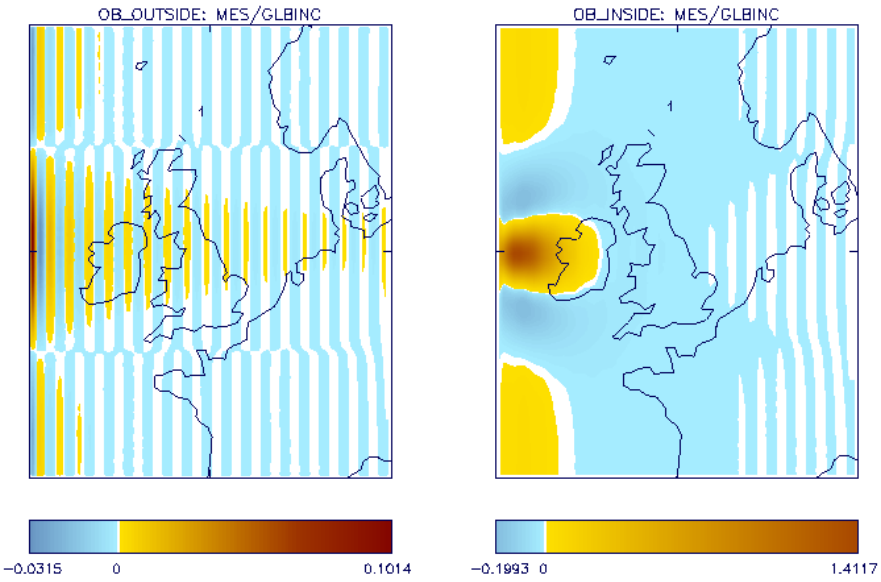

Fig. 2 U-wind analysis increments from a single U-wind observation of unit increment size at level 1 outside (left plot) and inside (right plot) the model domain. Global increments were used to construct a new first guess state. The results plotted here are relative to the background state

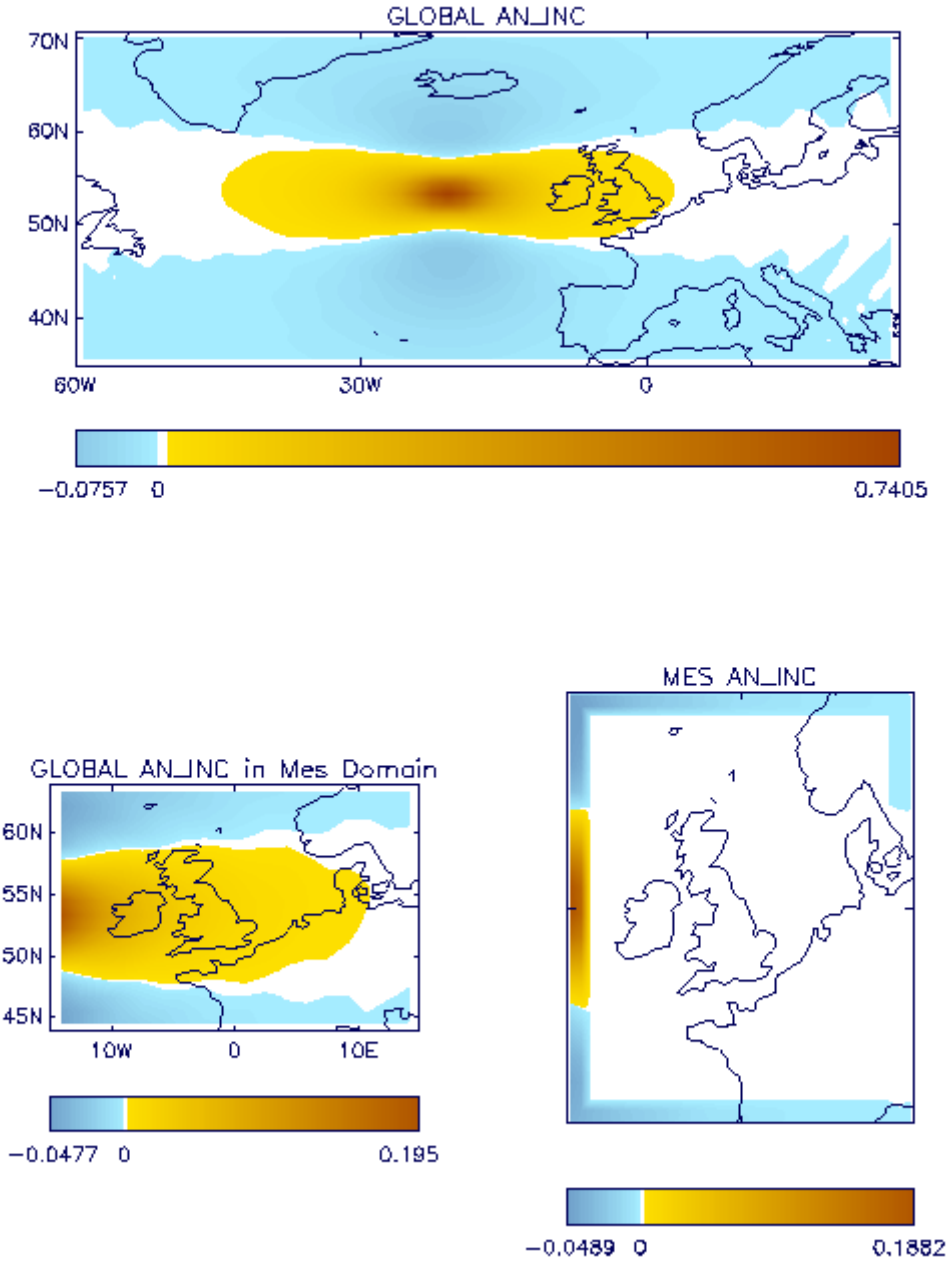

Fig. 3 U-wind analysis increments relative to the background from a single U-wind observation of unit increment size at level 1 outside the LAM domain. Global analysis increments over the wider LAM area (top plot), over the LAM domain (bottom left plot) and mesoscale analysis increments (bottom right plot). Boundary global increments were employed in the LAM

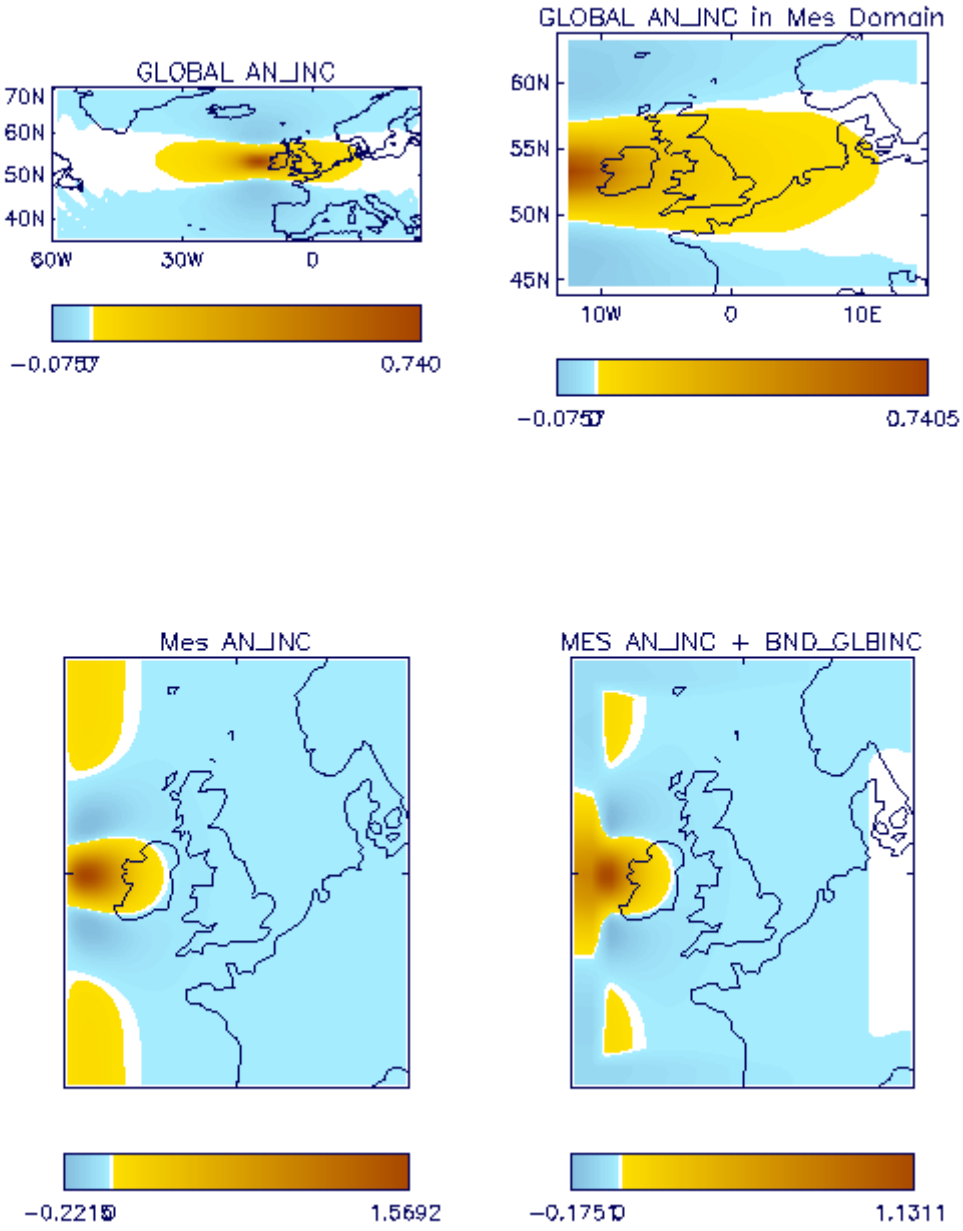

Fig. 4 U-wind analysis increments relative to the background from a single U-wind observation of unit increment size at level 1 inside the LAM domain. Global analysis increments over the wider LAM area (top left plot), over the LAM domain (top right plot) and mesoscale analysis increments without boundary global increments (bottom left plot) and with boundary global increments (bottom right plot)

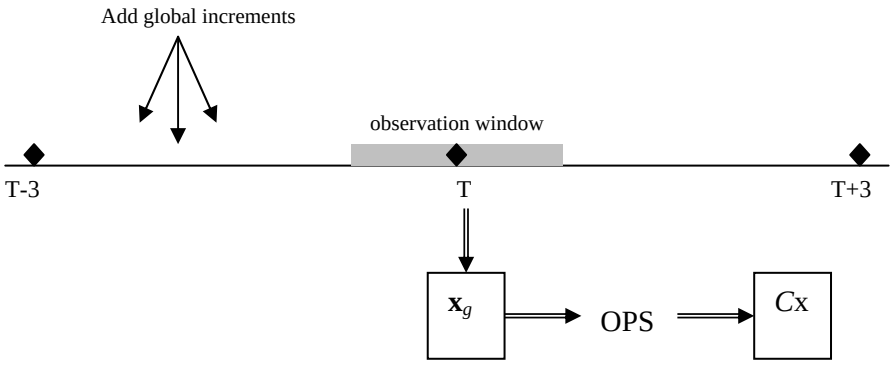

Fig. 5 An illustration of adding global increments with the IAU to produce the new first guess state, $\mathbf{x}_g$, at time T. OPS calculates the values of the first guess state at the location of the observations (Cx columns)

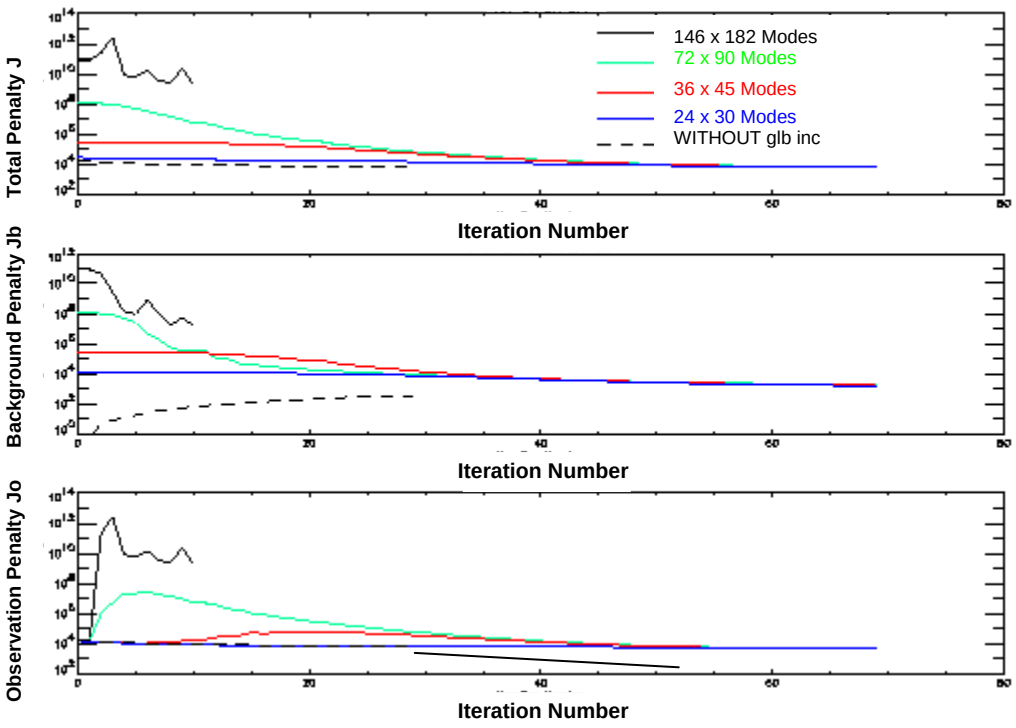

Fig. 6 The change of the total (top), background (middle) and observation (bottom) penalty function with iteration number for runs with the global increments (solid lines) and without (dashed line). Different number of modes were kept in the horizontal transforms and when all modes were kept in full resolution (146×182) the minimisation failed after 10 iterations (case QM12 – 19 Mar 2002)

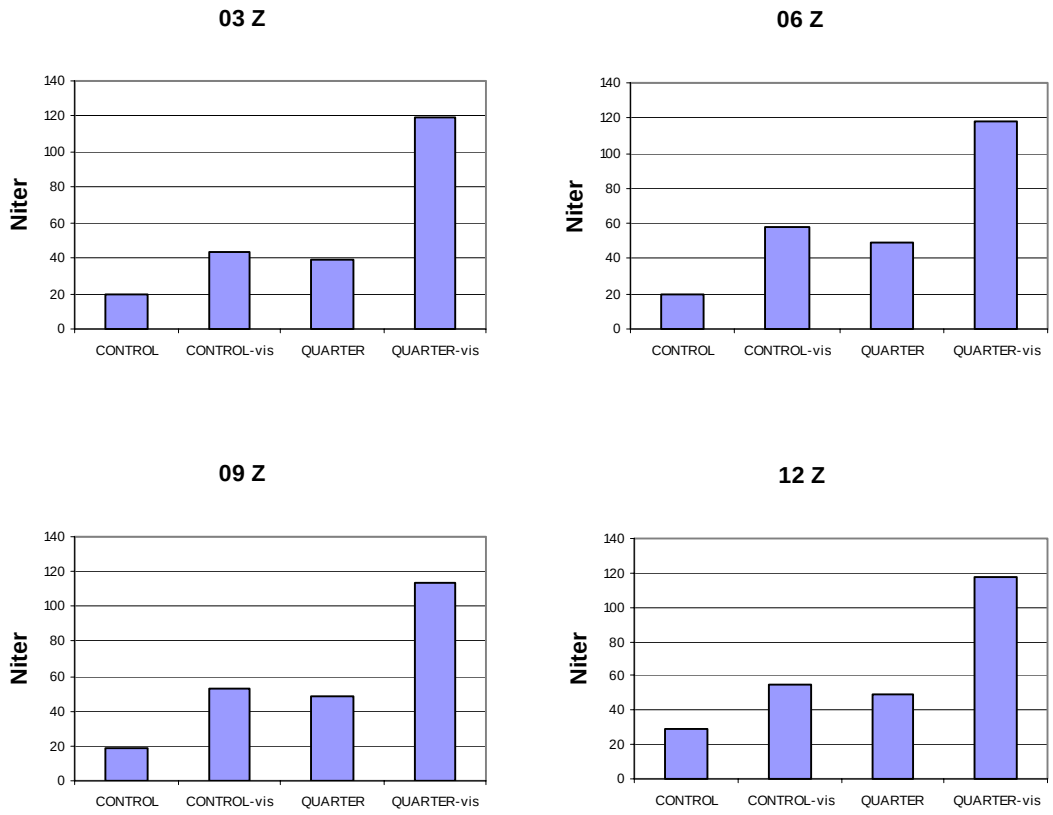

Fig. 7 The number of iterations in four analyses of a case study (29 Apr 2002) without (CONTROL) and with (GLBINC) the global increments and also without and with (vis) visibility assimilation

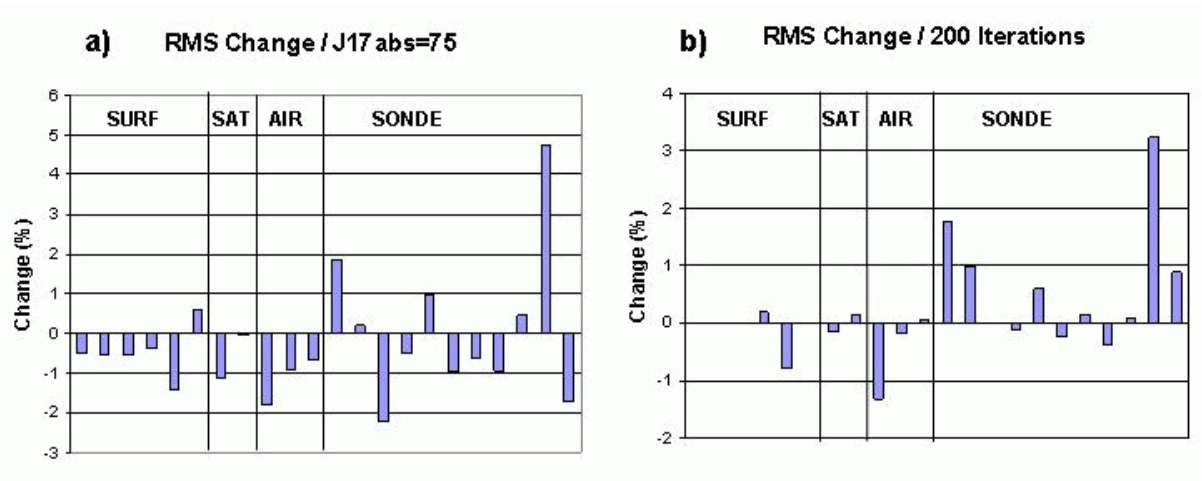

Fig. 8 Analysis fit to observations (case QM12 - 28 Aug 2003). The change in the RMS error for a run with the new first guess relative to a control run without global increments with a) the common convergence criterion (J17abs=75) and b) with 200 iterations. The bars correspond to different types of observations, i.e. surface (p, u, v, T, rh, vis), satellite (u, v), aircraft (θ, u, v) and sondes (θ at levels1, 15 and 25; u at the same levels; v at the same levels; rh at levels 1 and 15)

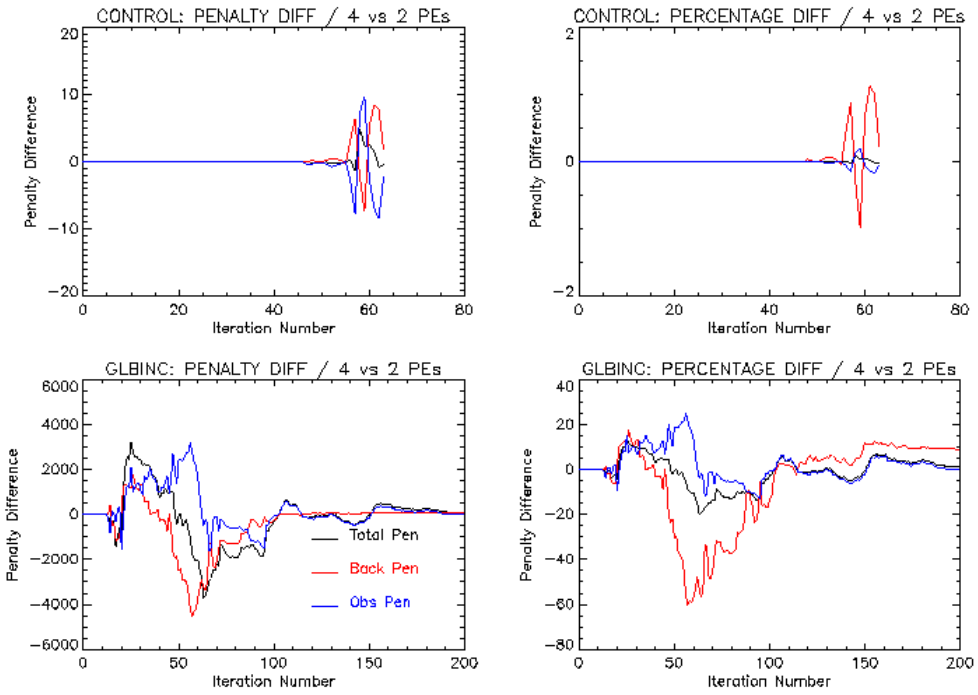

Fig. 9 The change in the penalty function with iteration number (plotted as difference and percentage difference) for a run with 4PEs relative to a run 2 PEs. The top plots are for a run without the global increments and the bottom for runs with the new first guess state. All runs were on the SX6 with Var Stable build 20.1. For the results plotted here, visibility was assimilated with the full non-linear operator. (Case: QM00 - 10 Dec 2003)

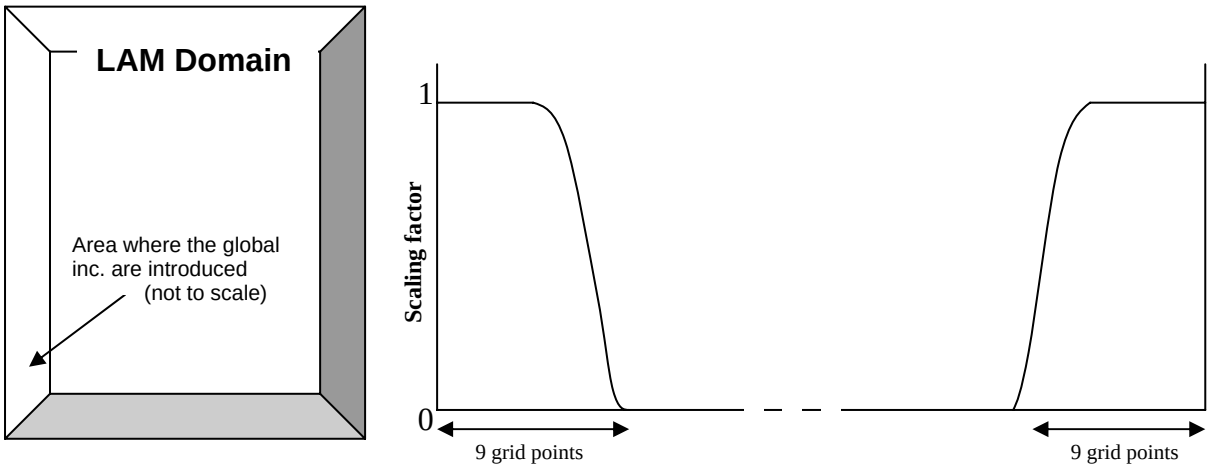

Fig. 10 A schematic illustration of the boundary increments scheme with the area where the global increments are blended with the analysis increments with a scaling factor along the x or y axis which decreases from 1 (5 points nearest the boundaries) to 0 (9 points away)

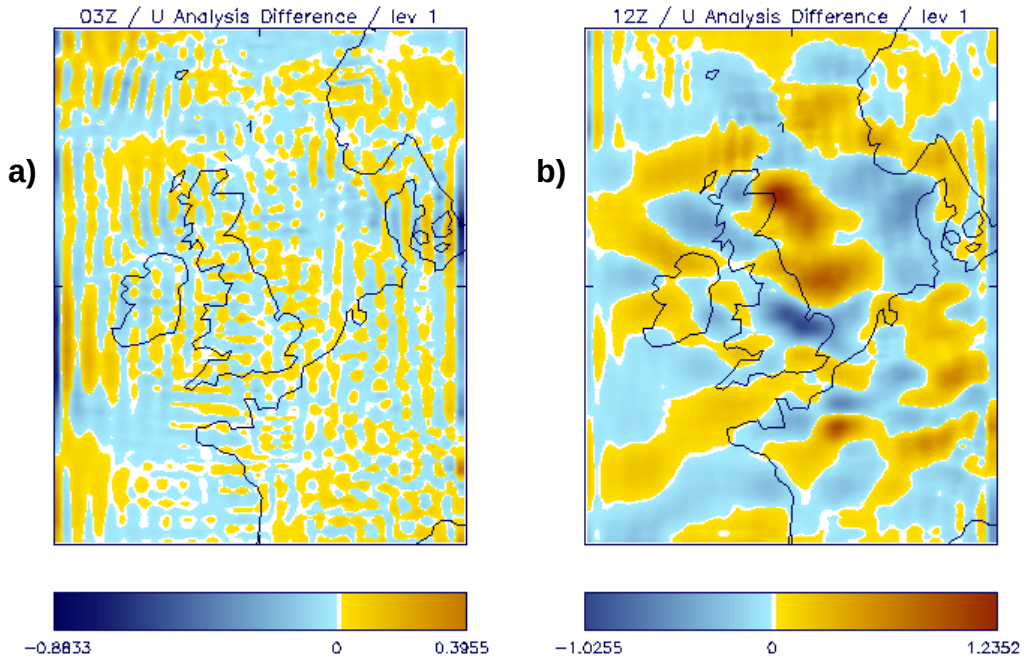

Fig. 11 U-wind analysis difference with and without the new first guess state for a case study (29April 2002) with 4 assimilation cycles at 03, 06, 09 and 12Z. The difference is plotted for a) the first cycle and b) the last cycle

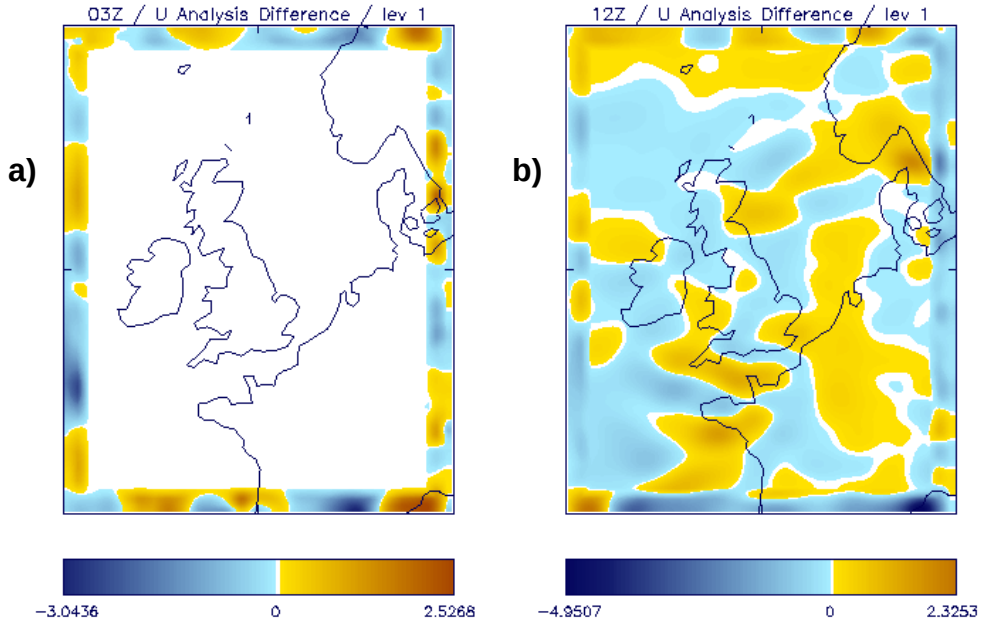

Fig. 12 U-wind analysis difference with and without the boundary global increments for a case study (01Oct 2002) with 4 assimilation cycles at 03, 06, 09 and 12Z. The difference is plotted for a) the first cycle and b) the last cycle

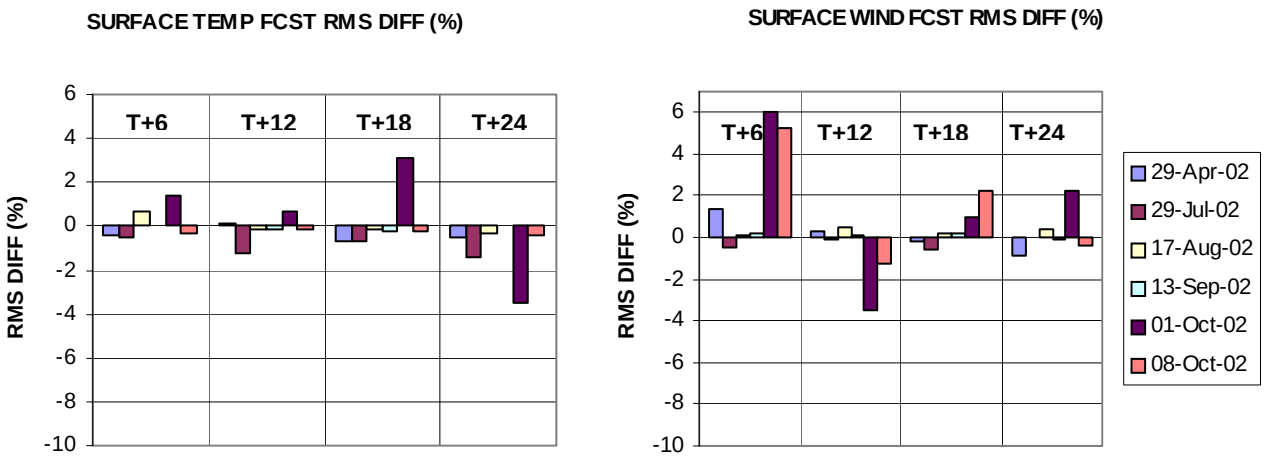

Fig. 13 Verification against observations for each of the 6 case studies. The change in the temperature and wind RMS error for test runs with the new first guess relative to control runs without.

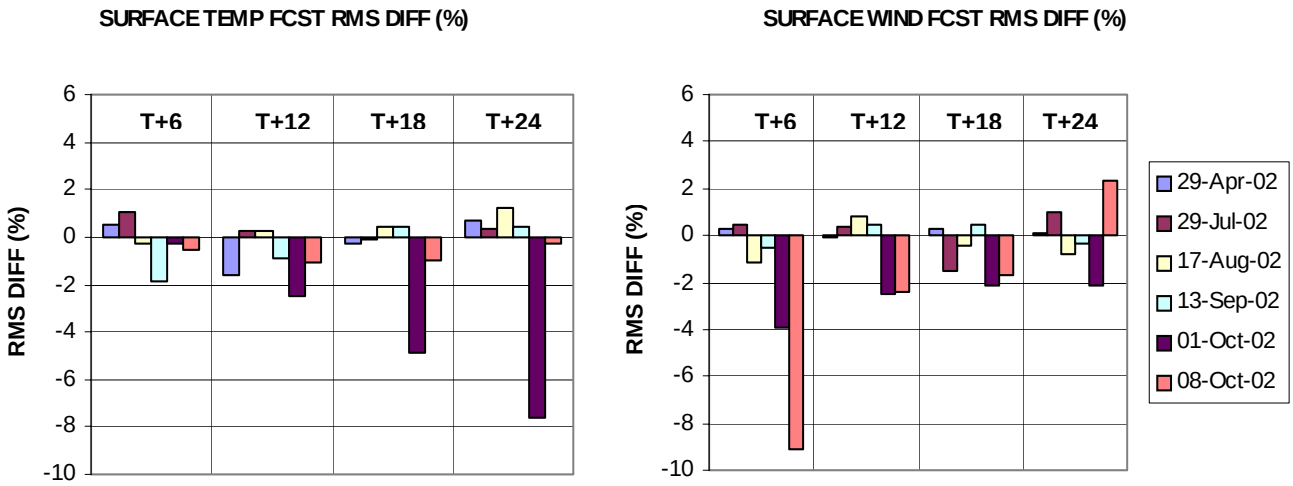

Fig.14 Same as Fig. 13 for test runs with the boundary increments as defined originally (global analysis minus global background)

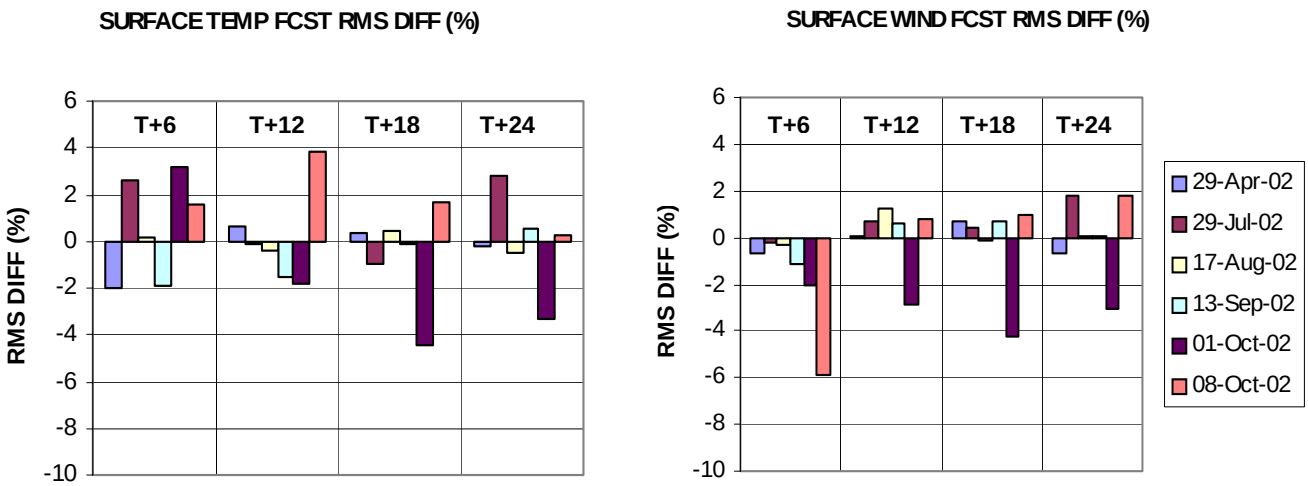

Fig. 15 Same as Fig. 13 for test runs with the boundary increments defined relative to the mesoscale background (global analysis minus mesoscale background)

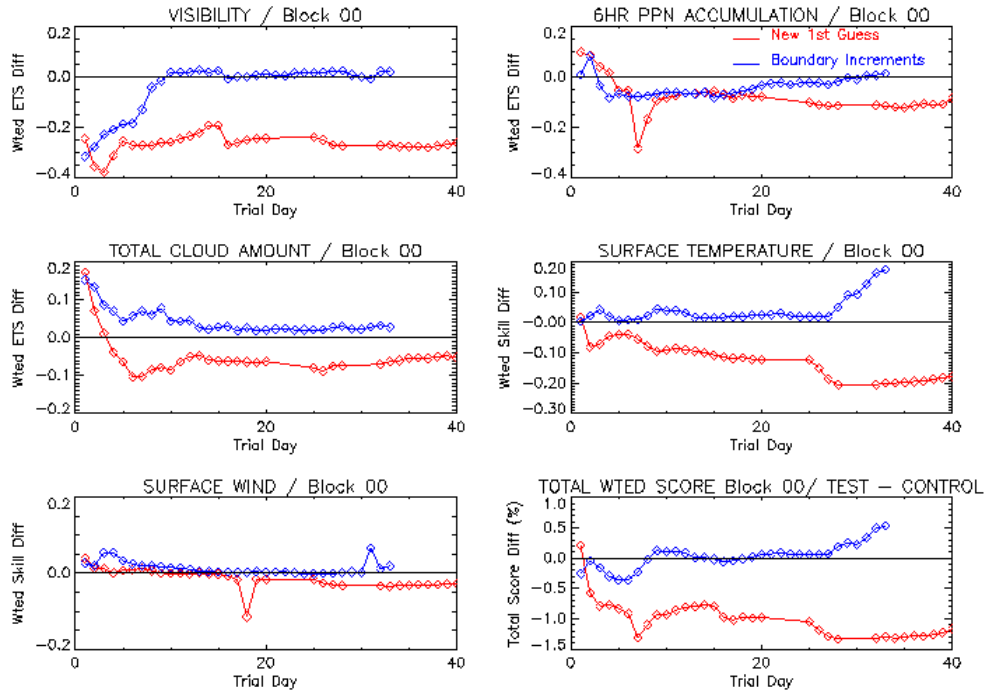

Fig. 16 The change in ETS and skill score for different parameters with the new first guess (red line) and the boundary increments (blue line) as a function of trial day. The trial periods for the two methods are different (see text). Results for block 00

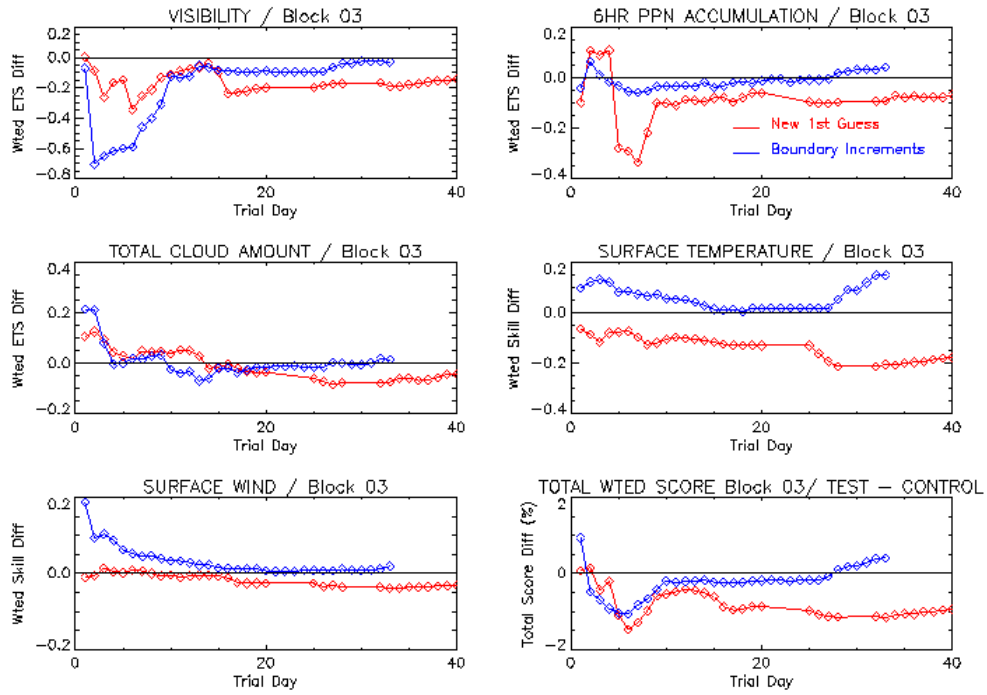

Fig. 17 Same as Fig. 16 for block 03

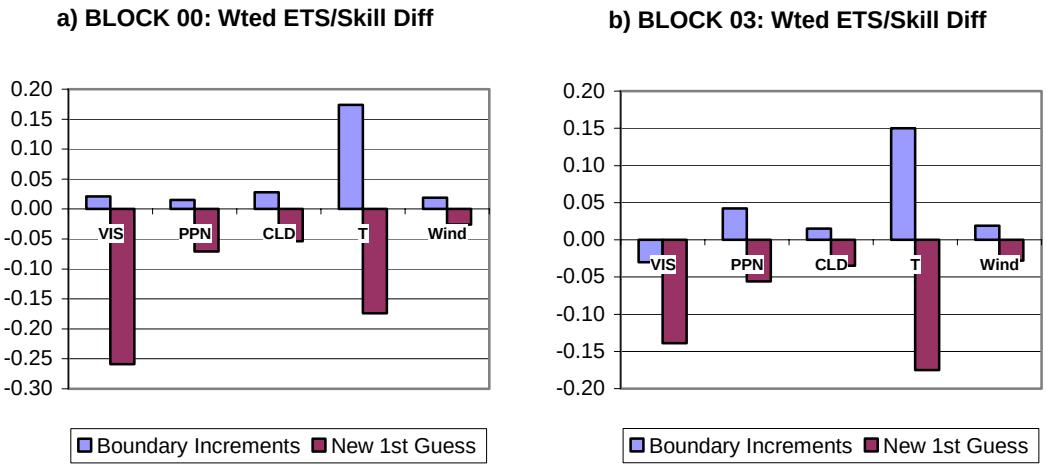

Fig. 18 Composite statistics from the trial runs, illustrating the change in ETS and skill score with the boundary global increments and the new first guess

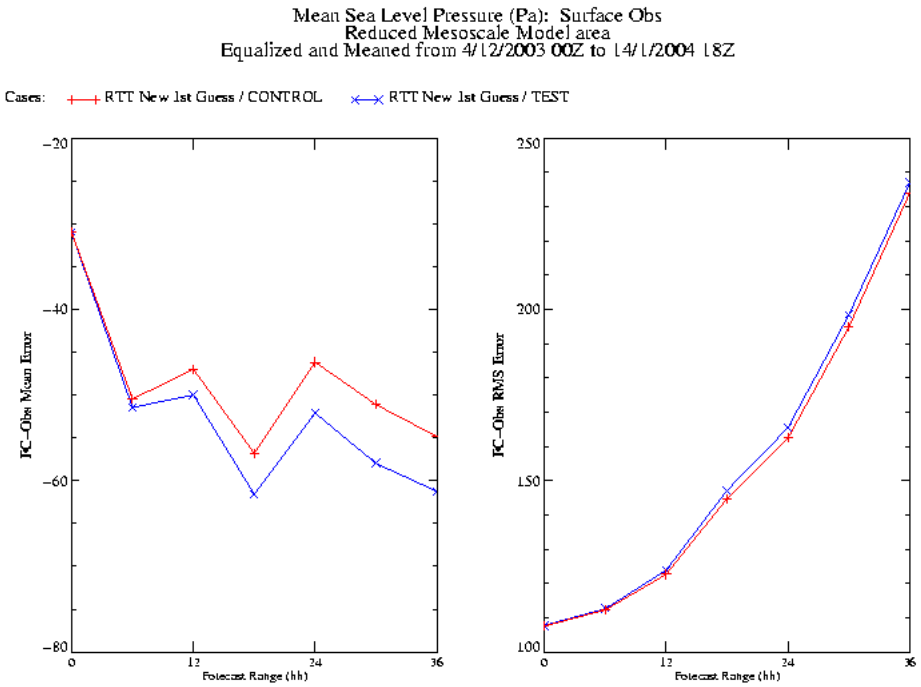

Fig. 19 The mean sea level pressure mean and rms error as a function of the forecast range with (blue) and without (red) the new first guess state for the composite of all the forecasts in the real time trial.

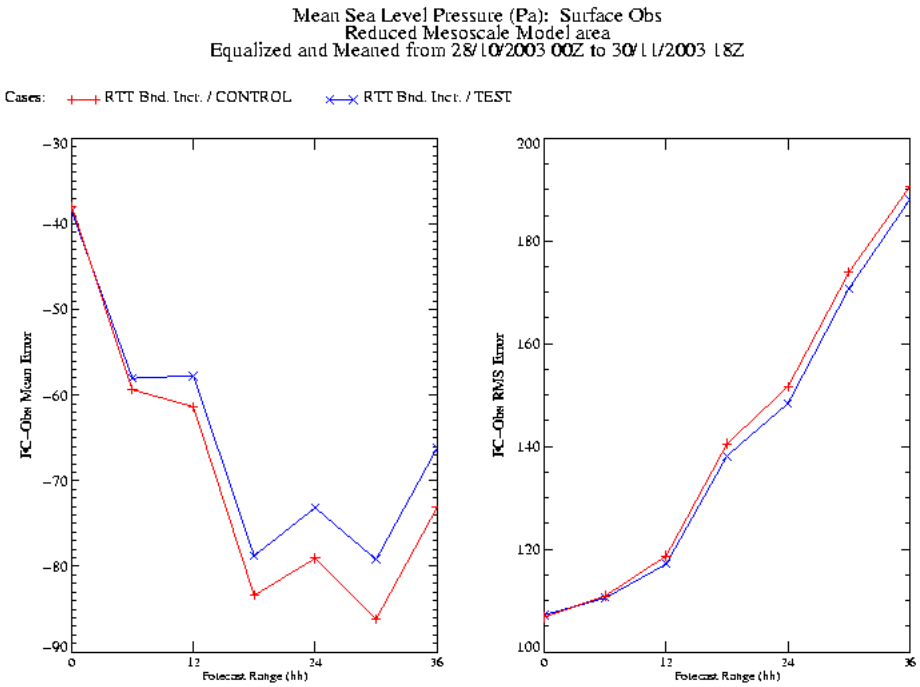

Fig. 20 The mean sea level pressure mean and rms error as a function of the forecast range with (blue) and without (red) the boundary global increments for the composite of all the forecasts in the real time trial.

References

Brozkova, R., D. Klaric, S. Ivatek-Sahdan, J.-F. Geleyn, V. Casse, M. Siroka, G. Radnoti, M. Janousek, K. Stadlbacher, and H. Seidl, DFI Blending: an alternative tool for preparation of the initial conditions for LAM, Research Activities in Atmospheric and Oceanic Modelling, Report No. 31, WMO/TD-No.1064, 2001.

Christidis, N., and B. Macpherson, Short-period Precipitation Forecasts: a Report from the Mesoscale Tiger Team Project, Forecasting Research Technical Report No. 436, 2004.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc, Unified notation for data assimilation: Operational, Sequential and Variational, *J. Met. Soc. Japan*, Special issue "Data Assimilation in Meteorology and Oceanography: Theory and Practice", *75*, No. 1B, 181-189, 1997.

Lorenc, A. C., S. P. Ballard, R. S. Bell, N. B. Ingleby, P. L. F. Andrews, D. M. Barker, J. R. Bray, A. M. Clayton, T. Dalby, D. Li, T. J. Payne, and F. W. Saunders, The Met Office Global 3-Dimensional Variational Data Assimilation Scheme, *Q. J. R. Meteorol. Soc.*, *126*, 2991-3012, 2000.

Lynch, P., Deducing the Wind from Vorticity and Divergence, *Mon. Wea. Rev.*, *116*, 86-93, 1988.

Siroka, M., G. Boloni, R. Brozkova, A. Dziedzic, C. Fischer, J.-F. Geleyn, A. Horanyi, W. Sadiki, and S. Soci, Innovative developments for a 3DVAR analysis in a limited area model: scale selection and blended cycling, Research Activities in Atmospheric and Oceanic Modelling, Report No. 31, WMO/TD-No.1064, 2001.

Wilson, C., S. Cusack, B. Chalcraft, M. Goeber, J. Bornemann, B. Macpherson, R. Renshaw, A. Maycock, and N. Robertson, Implementation of New Dynamics in the UK Mesoscale Model, Forecasting Research Technical Report No. 434, 2004.